Measuring students' thermal comfort and its impact on learning

Han Jiang
Worcester Polytechnic Institute
hjiang@wpi.edu

Matthew Iandoli
Worcester Polytechnic Institute
mjiandoli@wpi.edu

Steven Van Dessel
Worcester Polytechnic Institute
svandessel@wpi.edu

Shichao Liu
Worcester Polytechnic Institute
sliu8@wpi.edu

Jacob Whitehill
Worcester Polytechnic Institute
jrwhitehill@wpi.edu

ABSTRACT
Thermal comfort (TC) – how comfortable or satisfied a person is with the temperature of her/his surroundings – is one of the key factors influencing the *indoor environmental quality* of schools, libraries, and offices. We conducted an experiment to explore how TC can impact students' learning. University students ($n = 25$) were randomly assigned to different temperature conditions in an office environment ($25°C \rightarrow 30°C$, or $30°C \rightarrow 25°C$) that were implemented using a combination of heaters and air conditioners over a 1.25 hour session. The task of the participants was to learn from tutorial videos on three different topics, and a test was given after each tutorial. The results suggest that (1) changing the room temperature by a few degrees Celsius can stat. sig. impact students' self-reported TC; (2) the relationship between TC and learning exhibited an inverted U-curve, i.e., should be neither too uncomfortable nor too comfortable. We also explored different computer vision and sensor-based approaches to measure students' thermal comfort automatically. We found that (3) TC can be predicted automatically either from the room temperature or from an infra-red (IR) camera of the face; however, (4) TC prediction from a normal (visible-light) web camera is highly challenging, and only limited predictive power was found in the facial expression features to predict thermal comfort.

Keywords
thermal comfort, automated face analysis

1. INTRODUCTION
Most of the time that people learn takes place indoors. Primary and secondary school students are typically in school buildings for most of the day and do homework in their houses and apartments in the evenings. Adult learners may learn as part of their job in an office or pursue lifelong-learning opportunities at home. The *indoor environment quality* (IEQ) of where people learn, study, and work can have a significant impact on their physical well-being as well as their cognitive performance [1, 2].

The impact of IEQ on *learning* in particular has a special importance and has begun to interest architects, civil engineers, and educational psychologists in recent years [13]: Young learners in particular might be more sensitive to the influence of the environment due to their age or other physiological characteristics than adults. Students spend many hours each day in schools; however, since students typically have little control over their schools' physical environment, learners may feel great concern about their thermal comfort [8]. Thermal comfort (TC), which is a key component of IEQ, has been defined as "that condition of mind that expresses satisfaction with the thermal environment and is assessed by subjective evaluation" [3]. Prior work (see section below) has shown that suboptimal thermal comfort conditions can negatively affect students' learning. However, to our knowledge, no study to-date has explored the relationship between the impact of TC on learning and *time*. Is it possible that the effect of suboptimal TC could be mild during brief periods of learning but become more severe as the learning session continues? This is one of the questions we explore in this paper.

Measuring thermal comfort: Different people can experience the same temperature and environment differently, and just because one person has a high degree of thermal comfort does not mean her/his friend or peer will. Since thermal comfort is about a person's *satisfaction* with the thermal comfort, it depends not only on the environment itself, but also on the person's physiological and psychological *adaptability* [9, 7] to her/his environment. How adaptive a person is depends, in turn, on how and where a person grew up, e.g., her/his country of origin and its associated climate.

Due to the partially subjective nature of TC, most studies that sought to measure TC used questionnaires [12, 11, 9, 7]. While these are useful, they suffer from drawbacks such as (1) lack of temporal specificity, (2) recency/primacy effects, (3) disruption to regular activities. These can all lead to inaccurate measurements. Therefore, many researchers have explored alternative approaches based on various sensors (e.g., skin-based temperature sensors, cameras) to measure TC automatically [34, 23, 25, 22, 15, 17].

Han Jiang, Matthew Iandoli, Steven Van Dessel, Shichao Liu and Jacob Whitehill "Measuring students' thermal comfort and its impact on learning" In: *Proceedings of The 12th International Conference on Educational Data Mining (EDM 2019)*, Collin F. Lynch, Agathe Merceron, Michel Desmarais, & Roger Nkambou (eds.) 2019, pp. 89 - 98

Automatic facial expression recognition: One of the new forms of human observation that has been enabled by advances in machine learning and computer vision is based on automatic facial expression recognition. With technology, it is possible to automatically detect pain in the human body [18], student engagement [36], driver fatigue [33], and many other affective and cognitive states. Inspired by these studies, we explore in this paper whether automatic analysis of facial expression can help to detect a person's degree of thermal comfort.

Contributions: In our study, we (1) conduct a randomized experiment to explore the relationship between thermal comfort, the time-on-task, and learning. We also (2) explore different sensors and algorithmic approaches to estimating a person's thermal comfort automatically.

2. RELATED WORK

During the past 10 years there has been substantial interest (see [13, 26] for literature surveys) in measuring the impact of the IEQ on students' learning. In Table 1 we categorize the prior work on this subject in terms of IEQ factor (light, air, etc.) as well as the method of measuring learning (subjective impression (SI), test (T) performance, school scores (SS), and randomized experiment (RE)). In addition to studies specifically about thermal comfort (TC) [38], other factors of the IEQ such as lighting, air quality, and noise have been considered. Within this research domain, an important dimension of variability is how learning was measured – by asking participants their subjective impressions, from their school scores, or from a test conducted within the experiment itself. Another dimension of variability is whether the study was observational (i.e., compute a correlation between historical data of the IEQ and historical data of learning) or experimental (i.e., randomly assign participants to conditions). The latter is a generally considered to be the more powerful approach since it avoids many potential confounds (e.g., student engagement) and is the approach we pursue in our study.

2.1 Impact of TC on learning

[20, 8] used subjective impression as the learning performance. They both analyzed the relationships between the IEQ (light, air quality, thermal comfort and noise) and learning. [20] found that the learning performance was negatively correlated with the number of student complaints about IEQ. [8] also explored the students' satisfaction with IEQ, as well as the TC in particular, from survey data gathered from 631 university students. The results showed that satisfaction of IEQ of the classroom was related to the perceived effect of IEQ on learning. [27] conducted a 1-month test during May-June 2012 at a university in Romania. 18 students' test results of concentrated attention tests (Kraeplin test) and distributive attention test (Prague test)[31, 30] were recorded. The conductors used room temperature, relative humidity and CO_2 concentration to predict test scores. Their results suggested that these indoor environment factors could strongly impact students' learning performance. [35] conducted an experiment to explore the impact of air temperature on students' performance. The results indicated that with the same accuracy, students would increase their speed when performing the language-based and numerical performance tasks if the room temperature was reduced from $25°C$ to $20°C$ in late summer. [24] randomly assigned the participants into different conditions to perform a computer-based reading and learning task. They found that TC had a low and non-significant relationship with the performance; the participants in the extreme condition believed that the temperature had a larger negative impact on their performance than the participants in a normal condition. In [16], the researchers conducted an experiment to explore the impact of TC in 1-on-1 cognitive tasks when students are with a tutor. All the participants experienced all temperature conditions ($10°C$, $14°C$, $15°C$, $16°C$, $18°C$, $20°C$). Their experiment indicated that there was an inverted-U relationship between thermal sensation and pupils' learning performance. A seven point scale of thermal sensation, according to [3], was used. The meaning of the number from -3 to 3 was "cold", "cool", "slightly cool", "neutral", "slightly warm", "warm" and "hot" successfully. The results showed that students' performance was better in the cool or slightly cool conditions compared to the hot condition.

2.2 Measuring thermal comfort

How to measure thermal comfort has been explored for many years. While questionnaires from each person about her/his own TC is useful, they can be inconvenient and tedious. Researchers have thus sought to devise alternative measures that can be measured automatically from various sensors.

Environmental sensors: For instance, the PMV-PPD model, proposed by [12, 11], uses air temperature, mean radiant temperature, air velocity, humidity, and human variables to calculate the Predicted Mean Vote (PMV) of a group of people's averaged thermal sensation according to [3]. The Predicted Percentage of Dissatisfied (PPD) utilizes PMV to calculate the percentage of people who might complain about their thermal environment.

Body sensors: [34] used skin temperature sensors to collect upper extremity (finger, hand, forearm) skin temperatures and explored how these temperatures related to thermal sensation. [23] explored different configurations of where to place the temperature sensors on the body and identified particular configurations that were most effective.

Cameras: More recently, with the development of machine vision, researchers also explored predicting thermal comfort through cameras. [25] showed that the averaged forehead temperature from infrared (IR) images was correlated with people's thermal sensation and thermal comfort. [15, 17] leveraged the human thermoregulation process and then applied Eulerian Video Magnification algorithm[37] to filter the visible-light RGB images to predict thermoregulation states, which is one indicator of thermal comfort.

3. EXPERIMENT

In order to assess the impact of thermal comfort on learning and how this effect could change over time, we conducted a laboratory-based learning experiment (approved by WPI's IRB #18-0372) in which university students ($n = 25$) watched three lecture videos, answered surveys on their thermal comfort, and completed a quiz on what they learned. During the experiment, the indoor environment conditions were monitored and controlled according to a schedule defined by each participant's randomly assigned experimental

Table 1: **Related Work about the impact of indoor environment factors on learning. SI: subjective impression; T: test; SS: school scores; RE: randomized experiment**

	Light	Air	Thermal comfort	Noise	Other
SI	Lee, et al.[20] Choi, et al.[8] Marchand, et al.[24]	Kameda, et al.[19] Lee, et al.[20] Choi, et al.[8]	Lee, et al.[20] Choi, et al.[8] Marchand, et al.[24]	Lee, et al.[20] Choi, et al.[8] Marchand, et al.[24]	
T	Dorizas, et al.[10]	Kameda, et al.[19] Dorizas, et al.[10] Sarbu & Cristian.[27]	Dorizas, et al.[10]	Dorizas, et al.[10]	
SS		Haverinen-Shaughnessy, et al.[14]			Barrett, et al.[6] Barrett, et al.[5]
RE	Marchand, et al.[24]	Wargocki & David.[35]	Wargocki & David.[35] Marchand, et al.[24] Jiang, et al.[16]	Marchand, et al.[24]	

condition. We also deployed a variety of sensors – camera, environmental, and body – to measure the temperature of the environment and of each participant. These sensor measurements, along with participants' survey responses, allow us also to explore different automated approaches to estimating a person's thermal comfort.

3.1 Recruitment of participants

We recruited participants for the experiment through an email list at our university. In the end, 25 students (of whom 9 were female) participated in our experiment. All of them were either undergraduate or graduate students. Each participant was paid for \$20 gift card for his/her participation.

3.2 Procedure

This experiment was conducted on each participant individually and was divided into four sessions. Each session was 21 minutes. Therefore, every participant would sit at a desk around 84 minutes in total. In the first session (adaptation session), each participant gave informed consent, placed the skin-based temperature sensors on her/his body, and listened to the experimenter's instructions. The purpose of the adaptation session was to neutralize the potential impact of the outside weather conditions or physical activity (e.g., running to class) before the experiment. In each of the remaining three sessions, the participant watched a tutorial video (10 minutes), answered a quiz about it (<5 minutes), completed a thermal comfort survey (<5 minutes), and then took a break. The length of the break (21 min − VideoLength − QuizTime − SurveyTime) depended on how long the participant took to complete the quiz and survey. The order of the tutorial videos was randomized, as was the order of the temperature conditions (warm to neutral, or neutral to warm); see Conditions subsection below. Sensor measurements, including video of the face, were recorded throughout all three tutorial sessions.

After the participant finished putting on the body sensors, the experimenter started the videorecording from the laptop-based web camera, typed the participant's ID into the webpage, turned the time controller on, and then asked the participant to press the "Start" button whenever she/he was ready. The experimenter then left the room and stayed in the room next-door throughout the rest of the experiment. Using remote access software, the experimenter took an IR image of the participant at the beginning of each tutorial video during the tutorial sessions. See Figure 2 for a schematic of the procedure.

Figure 1: Experimental setup of the desk, laptop, and cameras.

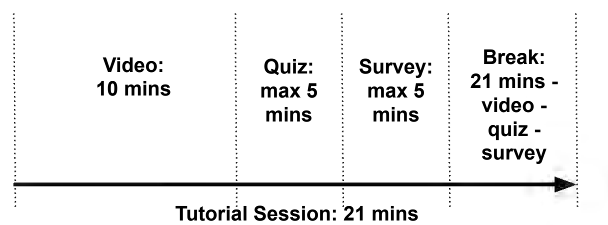

Figure 2: Tutorial session procedure

3.3 Environmental controls

We used 4 heaters (to increase the room temperature) and 1 air conditioner (to decrease temperature). In order to maintain the temperature at a constant level, we also deployed 3 thermal controllers. Moreover, in order to change the room temperature (from either warm to neutral, or neutral to warm), we also used 4 timers. To maintain the room temperature to be at least $25°C$, 1 heater was always turned on. 3 thermal controllers and 3 timers were connected to the other heaters. The thermal controllers were used to keep the room temperature around $30°C$. Timers were used to control when the heaters and air conditioners were turned on and off. The heaters and air conditioner were oriented so that the air did not blow directly onto the participant.

3.4 Sensors

All sensors were adjusted carefully before we started our experiment. They are listed as follows:

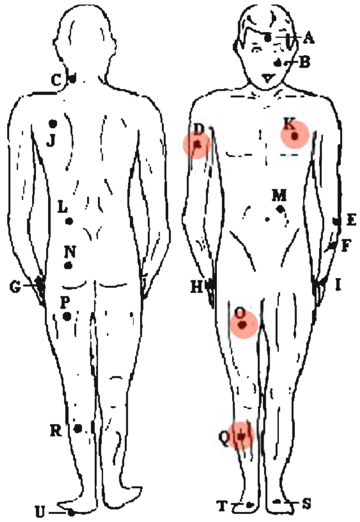

Figure 3: Positions of skin-based temperature sensors on the body.

1. 4 skin temperature sensors. We followed the positions in [23] (see Figure 3). These sensors were used to measure the participant's body temperatures at four different body locations and record the temperature every 1 minute. Sensors were attached using medical tape.

2. Room temperature sensors. These sensors were used to measure the room air temperature at different heights (0.1m, 0.6m, 1.1m and 1.7m) and recorded every 1 minute.

3. 1 web camera on the laptop pointed at the participant's face. Note that the video was lost for 1 out of 25 participants; hence, for our experiments on using the web camera to predict thermal comfort, $n = 24$.

4. 1 infrared (IR) camera pointed at the participant's face. The camera recorded only images, not video. Using the camera's temperature calibration software, the IR images can be used to estimate the participant's face temperature directly.

3.5 Materials

Tutorial videos: We used three 10 min-long tutorial videos and quizzes that were used in a prior study by [32]. The order in which the tutorial videos were presented to each participant was randomized; this was necessary to remove the potential confound that the subject matter, rather than the thermal comfort or time during the learning session, influenced the learning gains. All videos were about social, philosophical, and ethical issues: (1) honesty, (2) language and thought, and (3) empathy.

Thermal comfort survey: We used the same thermal comfort questionnaire survey as in [22, 21]. The survey asks questions such as, "Rate your whole body thermal sensation", "Rate your thermal body comfort", "How sleep/alert do you feel?", and "How easy/difficult is it to concentrate?" The scale was from -3 to +3 with a resolution of 0.1.

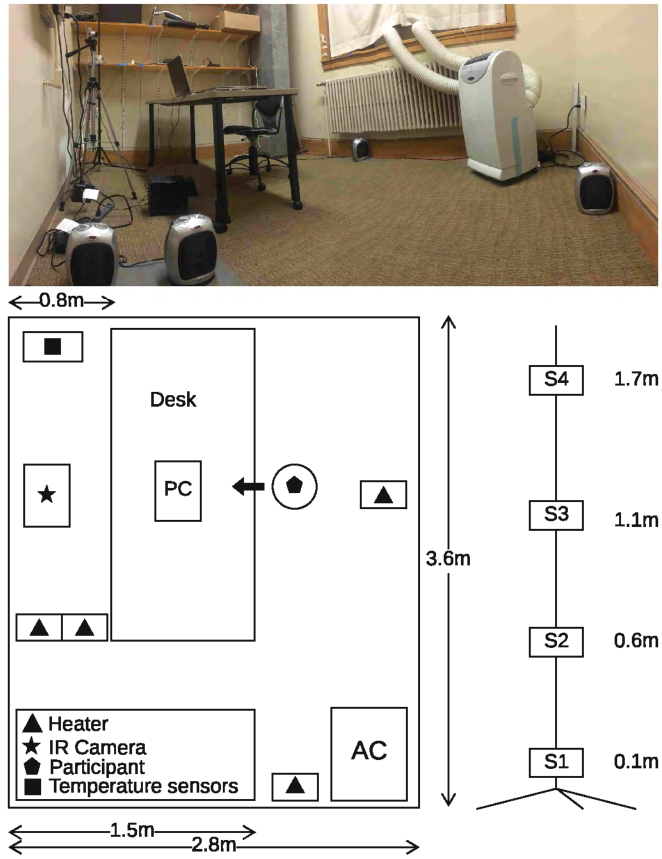

Figure 4: Top: Experiment lab Photo; Bottom Left: Top view of Lab and sensors' position. The participant was facing the direction with the arrow; Bottom Right: Room temperature sensors in different heights

3.6 Conditions

Each participant was randomly assigned to one of two temperature conditions: neutral to warm (25°C to 30°C), and warm to neutral (30°C to 25°C). By randomizing the thermal conditions, we avoid the potential confound that students' performance changed in different sessions not due to thermal comfort but due to other factors related to time, e.g., fatigue. If the participant was in the neutral to warm condition, the room temperature in the adaptation session was maintained at 25°C until the end of the first tutorial session; it was then increased to 30°C in the second tutorial session and was maintained at this level until the end of the third tutorial session. See Figure 5.

3.7 Data collection

Using the sensors, we collected several kinds of data from each person: (1) Video from the web-camera (at 30 fps); (2) Infrared images (1 every 21 minutes); (3) room temperature, CO_2, and relative humidity (1 measurement every minute); (4) body temperature (1 every minute for each sensor); (5) each participant's start/end times of each tutorial video, quiz, and survey; (6) each participant's quiz scores.

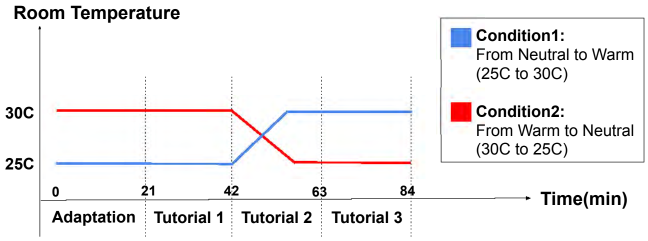

Figure 5: The change of room temperature in different conditions

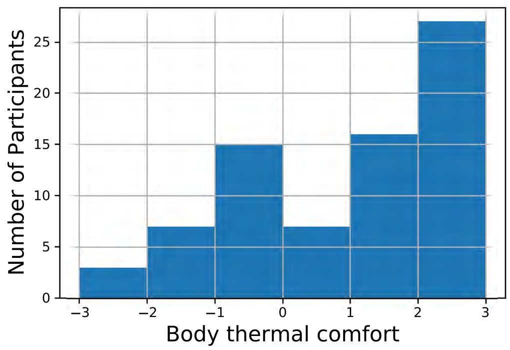

Figure 6: Histogram of thermal comfort in our experiment

4. ANALYSIS

Our analysis was focused on two questions: (1) what is the relationship between thermal comfort, temperature, and learning? (2) How can we use the various sensors to estimate participants' self-reported thermal comfort automatically?

4.1 Impact of room temperature on thermal comfort

In our experiment, the range of the room temperature was from 25°C to 30°C. This was not a huge change in the temperature. One of our goals was to assess whether this magnitude of temperature change could influence body thermal comfort. As defined in the thermal comfort survey that we used [3], the range of thermal comfort was from -3 to 3, where -3 means "very uncomfortable" and 3 means "very comfortable". Based on the histogram of body thermal comfort in our experiment in Figure 6, we see that the participants rarely (10 total votes) considered their thermal comfort to be highly uncomfortable (a rating of -3, -2). This indicated that our setting of the experiment was relatively comfortable for most of the participants. Did the modest temperature changes induced during the experiment impact participants' thermal comfort? To investigate, we considered models including either linear or quadratic terms for room temperature (computed as the average of the temperature sensors at different heights). The quadratic model did not give a stat. sig. better model fit, and hence we used a linear model; see Figure 7. The Pearson correlation between the model's predictions and self-reported thermal comfort scores was $r = -0.436$, $p < 0.001$, i.e., within the temperature range of our experiment, higher temperature resulted in lower thermal comfort. Based on the estimated regression coefficient, increasing the room temperature by one degree in our temperature range results in a reduction of thermal comfort by 0.32. Note that we also tried modeling thermal comfort and temperature (linearly) with a participant-specific offset as a random effect and obtained similar results.

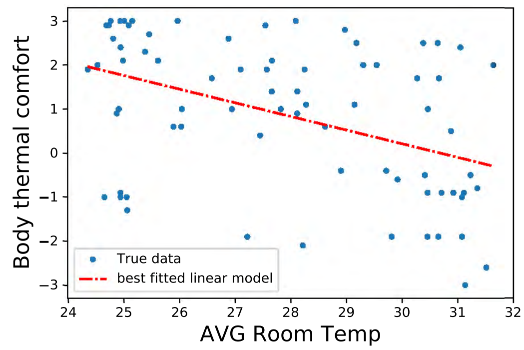

Figure 7: Thermal comfort VS Avg room temperature

4.2 Relationship between thermal comfort, learning, and time

After showing the change of room temperature in our experiment could influence the participants' thermal comfort, we assessed whether thermal comfort was related to participants' performance in the learning task. A scatter-plot of the quiz scores versus self-reported thermal comfort scores is shown in Figure 8. Neither the Pearson nor the Spearman correlations between quiz score and thermal comfort were significant. However, after visually examining the scatter-plot, we noticed a slight "inverted U" shape; this has also been noted in prior work [29, 28]. This shape indicates that when the participants felt too comfortable or too uncomfortable, their quiz score were lower; when the thermal comfort state was in the middle, their quiz score was higher. We found some support for this hypothesis in our data: the Spearman correlation between the *square* of self-reported thermal comfort and quiz score was negative ($r = -0.235$) and statistically significant ($p = 0.0042$). Tthe quadratic model of self-reported thermal comfort gives a stat. sig. better fit than the linear model (likelihood ratio test, $p = 0.002$).

To explore this more rigorously by accounting for repeated measures, we also used a mixed-effect model with a random effect to model an offset for each unique participant. Due to different tutorial videos having different difficulties, we also considered the video_id as the random effect. We studied the relationship between thermal comfort and quiz score within each of the three tutorial session (1, 2, 3) separately. To our surprise, in the first two tutorial session, the impact of the square of the body thermal comfort (i.e., TC^2) was not significant ($p > 0.05$). However, in the last (third) session, the impact was negative and stat. sig. ($p = 0.013$). The estimated magnitude was that a change in 1 level of thermal comfort decreases the quiz score by 0.2 points (the maximum score was 6 points). A possible interpretation is

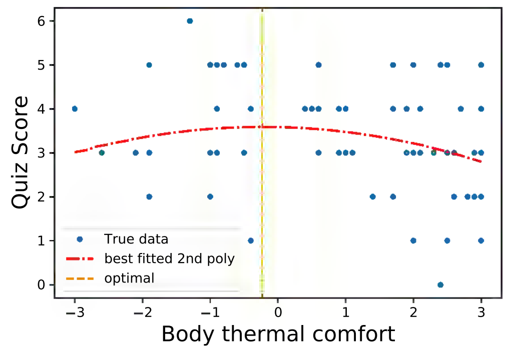

Figure 8: Thermal comfort VS Quiz score

Table 2: Effect size (Cohen's f^2) of TC^2 in each tutorial session

Session No.	Effect size
1	0.007
2	0.044
3	0.308

that, as time went on, the participants might feel more tired or bored. At first, they could force themselves to focus on the tutorial videos and answer questions. However, when they became fatigued or bored, an uncomfortable thermal comfort might start to show its influence. See Table 2 for the effect size(calculated based on the marginal R^2) in each tutorial session.

4.3 Relationship between thermal comfort and sleepiness

The survey that each participant completed after every tutorial session contained questions not just about thermal comfort, but also about how sleepy they felt. The values ranged from -3 (very sleepy) to +3 (very alert). The correlation between thermal comfort and sleepiness was positive (0.32) and stat. sig. ($p = 0.0084$).

4.4 Relationship between engagement and learning

To explore whether the perceived level of student engagement, as judged by an external observer, was related to students' learning, we manually labeled video frames from each participant's face video. We extracted 1 frame every 20 seconds for each of the 3 tutorial sessions of all the participants. These pictures were labeled for the appearance of 'engagement' following the definitions in [36]. Level 1 is "not engaged", level 2 is "nominally engaged", level 3 is "engaged", and level 4 is "very engaged"; see Figure 9 for a representative image of each label. During labeling, the images were randomized over time and also over participants; hence, the engagement scores were unbiased w.r.t. participants' self-reported thermal comfort. We averaged the engagement for each participant per each of the three tutorial sessions, and then used a mixed effect model to analyze the relationship between quiz score and engagement. The participant_id was still the random effect. Since we had a prior hypothesis

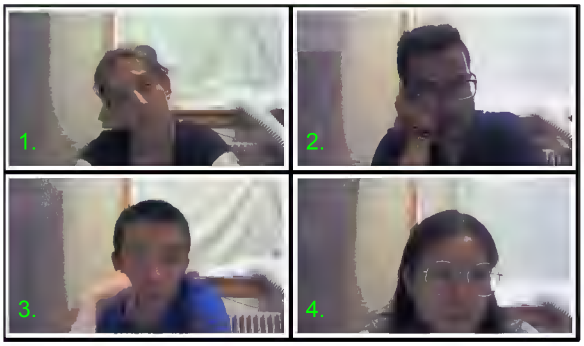

Figure 9: Participants in different engagement levels.

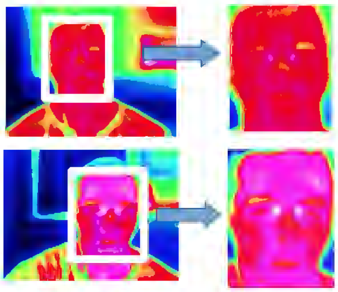

Figure 10: Manually cropped face for infrared images. Top: face when thermal comfort is -0.6. Bottom: face when thermal comfort is 2.7.

that engagement was positively correlated with learning, we used a 1-tailed t-test. The result showed that this positive correlation was significant ($p = 0.032$).

5. AUTOMATIC DETECTION OF THERMAL COMFORT

The primary method of estimating thermal comfort is via self-report on a survey. Might there be an automated way of obtaining this information that is less intrusive and gives higher temporal resolution? This could be useful to advance research on the IEQ and learning. Moreover, it could also set the stage for smart learning environments in which localized ventilation, heating, and cooling systems can optimize the thermal comfort for each learner. With these goals in mind, we explored several approaches to automatically estimating thermal comfort using the different sensors we deployed in our experiment.

5.1 Infrared camera

Per participant, 3 IR images were collected (one per tutorial session). From each IR image, we manually cropped the face for infrared images from IR camera and calculated the average face temperature for each tutorial session. For each IR image, we cropped the face between two ears for width, and from forehead to chin for length; see Figure 10. We then calculated the mean temperature within the face region and used it to predict thermal comfort. Using a mixed-effect model (with participant_id as a random effect), we found that the correlation between the face temperature, as computed from the calibrated IR image, and thermal comfort

Table 3: Skin Temp. VS Thermal comfort

Sensor	Pearson Correlation	p-value
D	**-0.273**	**0.018**
K	-0.174	0.136
O	-0.186	0.11
Q	**-0.28**	**0.015**

was $-.34$ ($p = 0.0029$). In other words, a hotter face was associated with lower thermal comfort.

5.2 Skin sensors of body temperature

We averaged the skin temperature from 4 skin sensors for each tutorial session. The correlations between thermal comfort and averaged skin temperature are shown in Table 3.

With statistical significance, the correlations of the skin temperature at position D and Q indicated that they had a negative correlation with body thermal comfort. These two correlations also remained significant when we applied the mixed-effect model and set participant_id as random effect.

5.3 Web camera

Even though the results of skin sensors and infrared cameras showed that we could use them to detect thermal comfort, we were still interested in whether an ordinary (visible light) web camera can be used to detect thermal comfort. In contrast to skin sensors, web cameras are less intrusive – they require no skin contact or medical tape. In contrast to IR cameras, they are less expensive and more widely available.

While one could consider a "black box" approach such as a CNN-LSTM in which all the pixels of an entire video segment is used to predict thermal comfort, the relatively small size of our dataset ($n = 24$) makes this approach difficult. Instead, we investigated whether the much lower-dimensional feature representation of facial expressions can reveal a person's thermal comfort. For example, we reported above that sleepiness is associated with thermal comfort, and this might be revealed in a person's facial expression; this approach was used in [33] to detect drowsiness when driving a car.

After watching the videos, our subjective impression was that predicting thermal comfort from the face was very difficult. In the temperature range of our experiment setting, the facial expressions in different temperature condition did not vary greatly. Nevertheless, we tried three approaches: (1) estimate thermal comfort directly from the average facial features values extracted from OpenFace [4] over the time series of face images; (2) estimate thermal comfort from a Gabor-filtered time series of facial features; and (3) train a recurrent neural network to analyze the raw time series.

5.3.1 *Individual face movements*

From each frame in each 10-minute video sequence just prior to the self-reported thermal comfort survey of each tutorial session of each participant, we used OpenFace to extract the facial action units (AUs 1, 2, 4, 5, 6, 7, 9, 10, 12, 14, 15, 17, 20, 23, 25, 26, 45). In addition, we also calculated the size of the face – this could be useful for determining if the participant leaned toward or away from the camera. Next, we extracted the head pose. Finally, we computed the distance between the eye-lids – this could give some measure of drowsiness.

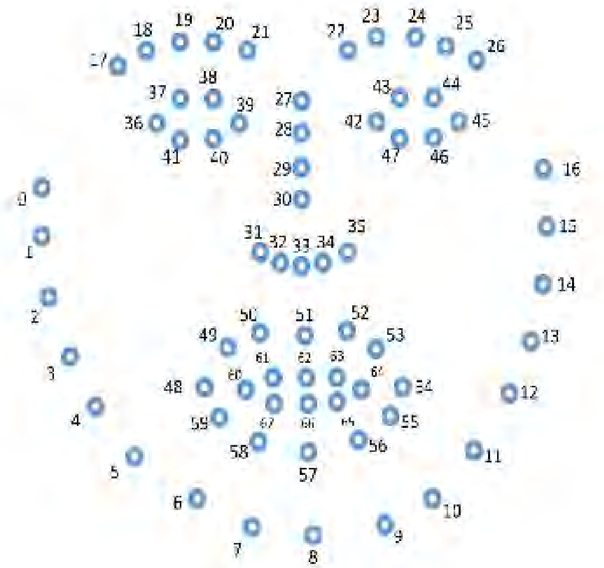

Figure 11: Landmarks from OpenFace

For the left eye, we first calculated the central point of landmark 37 and 38, the central point of landmark 41 and 40, and then, calculated the distance between the these two central points. For the right eye, we calculated the distance used landmark 43, 44, 47 and 46 as the same approach as the left eye. The eye-lid distance was the mean of the left distance and the right distance. We also estimated the size of the face box as an indication of whether a person was leaning towards or away from the camera: we first calculated the central of landmark 19 and 24, and then calculated the distance between the central and landmark 8, and also the distance between the landmark 0 and 16. The final face size was the product of the two distance. See Figure 11.

Using the above feature set, we examined the Pearson correlation between each mean feature value (averaged over each 10-minute time series) and self-reported thermal comfort. Only two features were stat. sig. correlated: AU 6 (Pearson $r = 0.244$, $p = 0.038$; see Figure 12) – cheek raiser – and the eye-lid distance, calculated by the landmarks on the eyes, was also correlated to thermal comfort with significant (Spearman $r = -0.27$, $p = 0.02$). The latter correlation suggests that smaller eye opening is associated with larger thermal comfort; this is consistent with the notion that thermal comfort that is "too high" may cause people to become sleepy.

5.3.2 *Gabor filtered time series*

A 1-D (temporal) Gabor filter is a complex-valued band-pass filter, with a specifiable center frequency and bandwidth, whose impulse response is local in both time and frequency; an example of the real component of one filter is shown in Figure 13. Gabor filters have been applied to var-

Figure 12: Example of AU 6 (`https://www.cs.cmu.edu/~face/facs.htm`

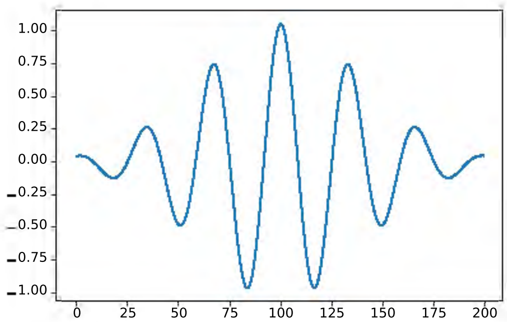

Figure 13: One example of real gabor filter. Frequency: 3.0; bandwidth: 0.9492

ious facial expression recognition tasks [33] and can capture certain patterns of a raw time series. For instance, they can capture wave-like patterns such as repeated blinking or eye closure. Here, we explored whether they could be helpful for predicting thermal comfort.

We applied Gabor filter to the AUs, face size, head pose, and eye-lid distance features. The frequency was selected from {8.0, 7.0, 6.0,5.0, 4.0, 3.0, 2.25, 1.6875, 1.2656, 0.9492, 0.7119, 0.5339, 0.4005, 0.3003, 0.2253, 0.1689, 0.01, 0.} and the bandwidth was selected from the same set without 0. Thus, 918 filters (the combination of 18 frequencies, 17 bandwidths and real, imaginary and energy Gabor filters) were applied to each AU and head features, which was the same filter bank as [33]. We used forward feature selection to pick the top 5 filtered features and then used linear regression on these top 5 features to predict thermal comfort. However, even the best Pearson correlation was very low ($r = 0.02$), suggesting that this approach had limited predictive power.

5.3.3 Recurrent neural networks

Recurrent neural networks such as LSTM and GRU, are powerful models for dealing with time series. We explored whether a GRU (Gated Recurrent Unit) network can analyze the facial expression series to estimate thermal comfort. We trained a GRU model from the feaures extracted using OpenFace described above using leave-one-person-out cross-validation to measure accuracy of the approach. Hyperparameters were selected from the sets {learning rate: {0.0001, 0.0005, 0.001}, hidden units: {8, 16, 32}, epoch: 50, optimizer: {Adam, SGD}. For each fold, we randomly selected 5 participants as the validation set (for hyperparameter validation), and the remaining 18 participants as the training set. Training every 5 epochs, the model would be applied to validation set and test set.

After tuning the hyper-parameters on the validation set, the best combination was {learning rate: 0.0005, hidden units: 32, epoch: 15, optimizer: Adam}. The average (over all 24 folds) correlation between predicted and actual thermal comfort scores was 0.248; the result was statistically significant ($p = 0.0425$, Wilcoxon signed-rank test). We note, however, that this result is no larger than the magnitude of the correlation between the eye-lid distance and thermal comfort reported above.

6. DISCUSSION AND CONCLUSION

We conducted an experiment in to investigate the relationship between thermal comfort and students' performance in a computer-based learning task in the classroom. We also explored different sensors and predictive models to measure thermal comfort automatically.

Key results: 1) Changing the room temperature by a few degrees Celsius could stat. sig. impact students' self-reported TC; (2) Our experimental data provide evidence that learning is optimal when thermal comfort is neither too high nor too low (inverted U relationship), corroborating prior work. However, we also found a more nuanced relationship than had been identified in prior literature: the impact of thermal comfort on learning was stronger during the third tutorial session (later in time) compared to the first two sessions. (3) Engagement, as labeled by an external observer, was correlated with learning. (4) Thermal comfort can be predicted from the face temperature using an IR camera. (5) Facial expression, at least in the ways we analyzed it, carries only limited information about thermal comfort.

Future work: Given a larger video dataset of face images and associated self-reported thermal comfort scores, we could explore more powerful prediction models that directly predict thermal comfort from the face pixels. This might offer more powerful information than the facial expression estimates from OpenFace.

Acknowledgements: This material is based upon work supported by the National Science Foundation under Grant Nos. #1551594 and #1822768.

7. REFERENCES

[1] Y. Al Horr, M. Arif, A. Kaushik, A. Mazroei, M. Katafygiotou, and E. Elsarrag. Occupant productivity and office indoor environment quality: A review of the literature. *Building and environment*, 105:369–389, 2016.

[2] M. Arif, M. Katafygiotou, A. Mazroei, A. Kaushik, E. Elsarrag, et al. Impact of indoor environmental quality on occupant well-being and comfort: A review of the literature. *International Journal of Sustainable Built Environment*, 5(1):1–11, 2016.

[3] ASHRAE. Standard 55-2004. thermal environmental conditions for human occupancy. *American Society of Heating, Refrigerating and Air-Conditioning Engineers*, 2004.

[4] T. Baltrušaitis, A. Zadeh, Y. Chong Lim, and L.-P. Morency. Openface 2.0: Facial behavior analysis toolkit. 2018.

[5] P. Barrett, F. Davies, Y. Zhang, and L. Barrett. The

impact of classroom design on pupils' learning: Final results of a holistic, multi-level analysis. *Building and Environment*, 89:118–133, 2015.

[6] P. Barrett, Y. Zhang, J. Moffat, and K. Kobbacy. A holistic, multi-level analysis identifying the impact of classroom design on pupils' learning. *Building and environment*, 59:678–689, 2013.

[7] G. S. Brager and R. J. De Dear. Thermal adaptation in the built environment: a literature review. *Energy and buildings*, 27(1):83–96, 1998.

[8] S. Choi, D. A. Guerin, H.-Y. Kim, J. K. Brigham, and T. Bauer. Indoor environmental quality of classrooms and student outcomes: A path analysis approach. *Journal of Learning Spaces*, 2(2):2013–2014, 2014.

[9] R. De Dear and G. S. Brager. Developing an adaptive model of thermal comfort and preference. 1998.

[10] P. V. Dorizas, M.-N. Assimakopoulos, and M. Santamouris. A holistic approach for the assessment of the indoor environmental quality, student productivity, and energy consumption in primary schools. *Environmental monitoring and assessment*, 187(5):259, 2015.

[11] P. Fanger. Moderate thermal environments determination of the pmv and ppd indices and specification of the conditions for thermal comfort. *ISO 7730*, 1984.

[12] P. O. Fanger et al. Thermal comfort. analysis and applications in environmental engineering. *Thermal comfort. Analysis and applications in environmental engineering.*, 1970.

[13] A. Gilavand. Investigating the impact of environmental factors on learning and academic achievement of elementary students. *Health Sciences*, 5(7S):360–369, 2016.

[14] U. Haverinen-Shaughnessy, D. Moschandreas, and R. Shaughnessy. Association between substandard classroom ventilation rates and students' academic achievement. *Indoor air*, 21(2):121–131, 2011.

[15] F. Jazizadeh and W. Jung. Personalized thermal comfort inference using rgb video images for distributed hvac control. *Applied Energy*, 220:829–841, 2018.

[16] J. Jiang, D. Wang, Y. Liu, Y. Xu, and J. Liu. A study on pupils' learning performance and thermal comfort of primary schools in china. *Building and Environment*, 134:102–113, 2018.

[17] W. Jung and F. Jazizadeh. Vision-based thermal comfort quantification for hvac control. *Building and Environment*, 2018.

[18] S. Kaltwang, O. Rudovic, and M. Pantic. Continuous pain intensity estimation from facial expressions. In *International Symposium on Visual Computing*, pages 368–377. Springer, 2012.

[19] K.-i. Kameda, S. Murakami, K. Ito, and T. Kaneko. Study on productivity in the classroom (part 3) nationwide questionnaire survey on the effects of ieq on learning. *Clima 2007 WellBeing Indoors*, 2006(Part 3), 2007.

[20] M. Lee, K. Mui, L. Wong, W. Chan, E. Lee, and C. Cheung. Student learning performance and indoor environmental quality (ieq) in air-conditioned university teaching rooms. *Building and Environment*, 49:238–244, 2012.

[21] A. Lipczynska, S. Schiavon, and L. T. Graham. Thermal comfort and self-reported productivity in an office with ceiling fans in the tropics. *Building and Environment*, 135:202–212, 2018.

[22] S. Liu, S. Schiavon, A. Kabanshi, and W. W. Nazaroff. Predicted percentage dissatisfied with ankle draft. *Indoor air*, 27(4):852–862, 2017.

[23] W. Liu, Z. Lian, Q. Deng, and Y. Liu. Evaluation of calculation methods of mean skin temperature for use in thermal comfort study. *Building and Environment*, 46(2):478–488, 2011.

[24] G. C. Marchand, N. M. Nardi, D. Reynolds, and S. Pamoukov. The impact of the classroom built environment on student perceptions and learning. *Journal of Environmental Psychology*, 40:187–197, 2014.

[25] B. Pavlin, G. Pernigotto, F. Cappelletti, P. Bison, R. Vidoni, and A. Gasparella. Real-time monitoring of occupants' thermal comfort through infrared imaging: A preliminary study. *Buildings*, 7(1):10, 2017.

[26] S. A. Samani and S. A. Samani. The impact of indoor lighting on students' learning performance in learning environments: A knowledge internalization perspective. *International Journal of Business and Social Science*, 3(24), 2012.

[27] I. Sarbu and C. Pacurar. Experimental and numerical research to assess indoor environment quality and schoolwork performance in university classrooms. *Building and Environment*, 93:141–154, 2015.

[28] O. Seppanen, W. J. Fisk, and Q. Lei. Room temperature and productivity in office work. 2006.

[29] O. A. Seppänen and W. Fisk. Some quantitative relations between indoor environmental quality and work performance or health. *Hvac&R Research*, 12(4):957–973, 2006.

[30] N. Srinivasan. Progress in brain research: Attention. 2009.

[31] V. Todea. Guide for psycho diagnosis laboratory. *Timisoara: Artpress Publishing House (in Romanian)*, 2008.

[32] S. Turkay and S. T. Moulton. The educational impact of whiteboard animations: An experiment using popular social science lessons. In *Proceedings of the 7th International Conference of Learning International Networks Consortium (LINC). Cambridge, MA, USA*, pages 283–91, 2016.

[33] E. Vural, M. Bartlett, G. Littlewort, M. Cetin, A. Ercil, and J. Movellan. Discrimination of moderate and acute drowsiness based on spontaneous facial expressions. In *2010 20th International Conference on Pattern Recognition*, pages 3874–3877. IEEE, 2010.

[34] D. Wang, H. Zhang, E. Arens, and C. Huizenga. Observations of upper-extremity skin temperature and corresponding overall-body thermal sensations and comfort. *Building and Environment*, 42(12):3933–3943, 2007.

[35] P. Wargocki and D. P. Wyon. The effects of moderately raised classroom temperatures and classroom ventilation rate on the performance of schoolwork by children (rp-1257). *Hvac&R Research*, 13(2):193–220, 2007.

[36] J. Whitehill, Z. Serpell, Y.-C. Lin, A. Foster, and J. R. Movellan. The faces of engagement: Automatic recognition of student engagementfrom facial expressions. *IEEE Transactions on Affective Computing*, 5(1):86–98, 2014.

[37] H.-Y. Wu, M. Rubinstein, E. Shih, J. Guttag, F. Durand, and W. Freeman. Eulerian video magnification for revealing subtle changes in the world. 2012.

[38] Z. S. Zomorodian, M. Tahsildoost, and M. Hafezi. Thermal comfort in educational buildings: A review article. *Renewable and sustainable energy reviews*, 59:895–906, 2016.